NTT DATA
Trusted Global Innovator

NEWS RELEASE

August 29, 2019

NTT DATA announces
Apache Kafka based end to end large-scale data processing platform solution
--Easily develop production systems to process real-time data such as sensor data or location data--

NTT DATA Corporation

Tokyo, Japan, - August 29 - NTT DATA Corporation (NTT DATA) announced the immediate availability of its new "Kafka Solution", which enables real-time processing of large-scale data.

Apache Kafka is an open-source distributed messaging platform (note 1), which is capable of handling large-scale data generated by sources such as IoT devices. Wide use of such devices in recent years made it difficult to gather all the data efficiently because it was required to design and develop data ingestion mechanisms per system. Apache Kafka solves this problem by providing a common messaging platform, which makes real-time processing easier while providing advanced features and high flexibility. Therefore, Apache Kafka is being used in many real-world scenarios such as congestion-aware route guidance based on GPS data and speed sensor data, real-time fraud detection in financial transactions etc.

The "Kafka Solution" offered by NTT DATA covers all the phases of analytics platform creation from consultation and PoC to platform development and operations. Customers who are considering real-time processing of large-scale data can take full advantage of this solution to create scalable and reliable production quality environments, which are made for long-term operation.

NTT DATA will aim for a total sales of 10 billion yen over three years through this solution.

[Background]

In recent years, real-time processing of large-scale data and taking actions within seconds to minutes based on the results has been under the spotlight. Apache Kafka is an open-source distributed messaging platform which delivers large-scale data generated by sensors and other mediums to real-time processing platforms. Although Apache Kafka is a convenient product that incorporates various functions and designed with commercial use in mind, it raises the hurdle for using in production systems. This is because it is a relatively new product with a sparse track record and requires knowledge of distribution systems.

Prior to this, NTT DATA has also offered "Hadoop/Spark Solution" (Note 2), which allows customers to develop Apache Hadoop (Note 3) or Apache Spark (Note 4) platforms. Apache Hadoop and Apache Hadoop are open-source large-scale data processing frameworks. Prior to announcing "Kafka Solution", NTT DATA has conducted original testing of Apache Kafka and accumulated knowledge to use and operate Apache Kafka in commercial systems with confidence to meet the rising needs toward real-time application of large-scale data. Taking advantage of the knowledge that has been accumulated, NTT DATA has decided to start offering "Kafka Solution" to support a wide range of real-time processing platforms for large-scale data.

[Features]

"Kafka Solution" offers a wide range of technical support from consultation, PoC, infrastructure development, and operation of real-time processing platforms for large-scale data using Apache Kafka. Utilizing accumulated knowledge and unique technologies, NTT DATA provides real-time processing platform that are excellent in scalability, reliability, and operability for customers who are considering real-time processing of large-scale data.

本件に関するお問い合わせ先　株式会社NTTデータ　広報部／TEL.03-5546-8051　〒135-6033　東京都江東区豊洲3-3-3　豊洲センタービル

■ A highly reliable distributed messaging platform with fewer servers

NTT DATA's solution offers a platform with high reliability and operability with fewer number of servers compared to standard Apache Kafka. They have achieved this by developing custom built fault-tolerance mechanisms, which have been independently verified, in addition to the standard fault-tolerance mechanisms of Apache Kafka.

■ One-stop consultation and integration of the entire data analytics platform according to the customer's environment and requirements

NTT DATA will provide consultation services for customers who are considering to use Apache Kafka. They will help to design the overall analytics platform architecture with customer's system requirements and future expandability in mind. Furthermore, by combining with "Hadoop/Spark Solution", they will support integration of the entire data analytics platform including batch processing platform for large-scale data.

■ Support services by professional engineers

NTT DATA will answer inquiries regarding troubleshooting and other questions related to creation and operation of systems that use Apache Kafka. Their support service can be used in combination with the existing Hadoop/Spark support service. All inquiries are answered by professional engineers who can analyze source code if necessary. They will provide continuous support services for systems that operate for longer period of time where upgrading is difficult.

[Use case examples]

The real-time data processing platform offered by this solution can be used for various purposes. For example, information gathered via GPS or speed sensors can be used to calculate the traffic volume to conduct real-time route guidance. Another example is gathering and analyzing transaction information from financial institutes and credit cards for real-time fraud detection to prevent losses or reduce damages.

For such systems, "Kafka Solution" ensures that Apache Kafka delivers and aggregates data sent from a large number of sensors and devices. It also provides total support from the design and construction of the infrastructure considering the operability, the future data volume increase, and the expandability that can withstand changes in the processing content, and the long-term operation after the service starts.

[Future developments]

NTT DATA will promote the use of systems that use Apache Kafka, regardless of field or industry, and aim for a total sales of 10 billion yen over three years.

NTT DATA will aim for a total sales of 10 billion yen over three years, by promoting systems incorporating Apache Kafka regardless of industry or field. At the moment, NTT DATA has over 300 professional engineers who can support Apache Hadoop/Apache Spark. These engineers will also be supporting Apache Kafka, and the support system is scheduled for further developments as well. Furthermore, initiatives for further improvements in reliability and operability for use in commercial systems, along with contributions to the development community are also to be continued.

(Note 1) In here, the platform for message (data) exchange between multiple systems is called a messaging platform. A distributed messaging system is a messaging platform created by multiple servers.

(Note 2) Hadoop/Spark Solution https://www.nttdata.com/jp/ja/lineup/bizxaas_hadoop/

(Note 3) Apache Hadoop http://hadoop.apache.org/

(Note 4) Apache Spark http://spark.apache.org/

■For more information, please contact:
Masaru Dobashi, Kenichi Shida, Toru Sasaki
Technology and Innovation General Headquarters
Systems Engineering Headquarters
Production Engineering Department
OSS Professional Services
NTT DATA Corporation
hadoop@kits.nttdata.co.jp

[Reference]

■Use cases

Use cases of Apache Kafka by NTT DATA are introduced at events such as the following overseas conferences:

https://kafka-summit.org/sessions/worldwide-scalable-resilient-messaging-services-kafka-kafka-streams/

https://conferences.oreilly.com/strata/strata-ny-2018/public/schedule/detail/69164

https://www.slideshare.net/hadoopxnttdata/hadoop-spark-conference-japan-2019nttdata (Japanese only)